

Heterogeneous Architecture Support for Wireless Network Dynamics and Mobility

JOSEPH P. MACKER

VINCENT D. PARK

Protocol Engineering and Advanced Networking (Protean) Research Group

Communication Systems Branch

Information Technology Division

November XX, 2000

Purpose

In this paper, we examine network architecture issues and protocols for applicability to mobile and dynamic wireless communication networks. Our focus is on networking technology to provide heterogeneous support and interoperability with existing and planned internetworking infrastructures and applications. First, we discuss engineering issues and technology components across a broad range of mobile networking system elements. Second, we focus in detail on a number of critical technology areas relating to mobility support and interoperability with the Internet Protocol (IP) suite. Our goal is to highlight important issues, provide guidance in examining architectural and protocol tradeoffs and help assess the maturity of technology components for near term deployment of network systems and further experimentation in ongoing mobile network projects. Since there is not a definitive "one size fits all" solution space to the design components involved in mobile networking, a future follow-on study is planned to provide more project-specific recommendations for specific wireless networking technology architectures and requirements—e.g., Interoperable Networks for Secure Communications (INSC).

Approach

Our approach in this paper involves discussion and analysis of several existing and emerging networking technology areas in the context of wireless tactical networks. While there are numerous commercial and Internet standard development activities relating to wireless networking protocols, we assume, that the goal architecture of interest must satisfy the challenging operational conditions and heterogeneous nature of present and future Navy and Marine Corps networking media. In our discussion, we target this specific application area of interest.

Future enhancements in military wireless networking capability require technical innovation and design improvements at and across many layers of the network communications stack. Some needed wireless improvements in link bandwidth and local link resource sharing may be realized by developing lower layer technology enhancements such as coding and modulation advancements, active antennas, and innovative, shared media access protocols. While important areas for continued research and development, the authors contend that innovations in lower layer technologies alone will not realize the network-centric warfighting capability that is being planned for and envisioned in future DoD systems. Application, data transport, and internetwork service layers (e.g., heterogeneous system routing, mobility services) are key technology areas for additional focus.

Wireless environmental features present specific, new challenges to operational military networks. Furthermore, when node mobility or potential network topology dynamics are expected the technical challenges and architectural tradeoff issues increase in complexity. An important focus of this report involves examining the interaction, boundary conditions, and potential performance tradeoffs of protocols and wireless networking subsystems. This involves considering partial or extensive dynamics and mobility of both end users and the infrastructure itself. We examine these technical challenges and the associated design tradeoffs and choices,

but we focus mainly on technology components and system tradeoffs at the internetwork layer and above.

Problem

The envisioned military networking environments for future wireless technology push requirement limits of scalability, robustness, adaptability and efficiency for network protocols and applications. Network protocols and applications available in commercial products today are typically designed for operation in quasi-static hardwired networks. Protocols and applications for future military networks must be designed for efficient operation under the resource constraints and behavioral dynamics of wireless networks. While this topic has attracted renewed research interest in recent years, the focus has been limited to a few isolated "building blocks"—e.g., developing unicast routing protocols for mobile ad hoc networks or optimizing the Transport Control Protocol (TCP) for use over satellite links. Synergistic, interoperable solutions that provide a more unified approach to networking in a heterogeneous wireless infrastructure and offer a near equivalent capability to that currently available in hardwired networks are far from complete. To reiterate, the commercial technology has yet to meet the public expectations of today; thus, realizing the DoD expectations of future military networks is even more challenging. Since mobile networking technology continues to evolve rapidly along with Internet-based networking technology, any envisioned networking software and hardware system should be designed with *extensibility* and *modularity* up-front. This greatly increases the potential advantage of future "economies of scale" and interoperability with ongoing advances.

A complete mobile networking architecture involves many subcomponents and we examine and discuss a number of these different areas in subsequent sections. One significant difference in emphasis between present commercial requirements for mobile systems and military requirements is in the need for significant infrastructure mobility support. Commercial mobile networking systems typically assume a static or quasi-static backbone infrastructure with end users that change their point of attachment as they migrate (i.e., most of the mobility resides in end users, while the network infrastructure and attachment points remain fixed). In military applications, wireless network infrastructure nodes (e.g., routers) are often on the move in addition to or in conjunction with end users. Thus, the infrastructure nodes also require adaptability and some degree of autoconfiguration. We examine this aspect of mobile networking technology and the alternative architectures and technologies available. It is rather sensible to imagine these alternative technologies working together at a broader internetworking level to solve different scenario requirements. Figure 1 demonstrates a set of high-level requirement considerations for developing a suitable mobile network architecture design. The authors believe that future DoD wireless tactical networks require system support across a hybrid set of requirements as shown in Figure 1.

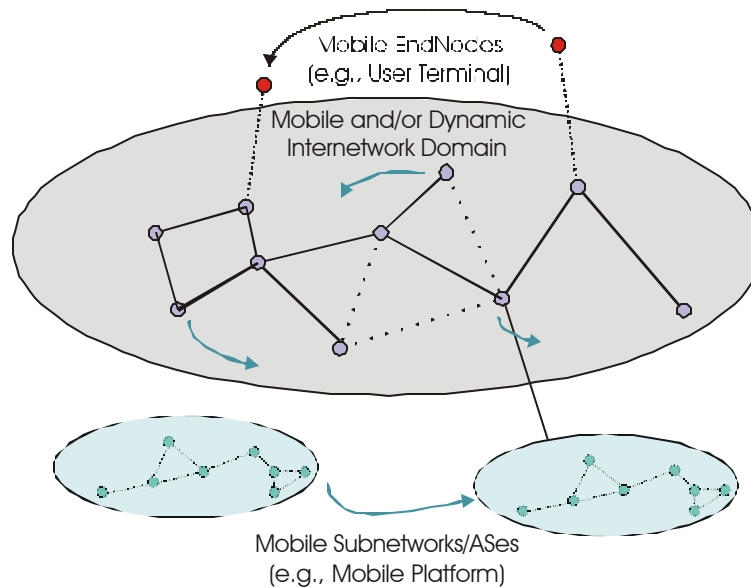


Figure 1: Dynamic and/or mobile internetwork architecture considerations

Mobile, Wireless Networks vs. Fixed Networks

There are numerous operational factors that distinguish mobile, wireless networks from fixed networks. Some of these include the following:

- nominally lower bandwidth network links (range, power, spectrum, and antenna tradeoffs)
- unique medium access channel (MAC) technologies (e.g., collision avoidance)
- semi-broadcast nature of some wireless multiple access media (hidden terminals)
- increased likelihood of interference and congestion (e.g., due to bandwidth constraints, frequency restrictions or channel access techniques)
- more frequent topological changes (e.g., due to node mobility, channel propagation effects, failures, power control, or antenna dynamics)
- higher loss rates (e.g., due to interference, congestion or dynamics)
- potentially higher delays and jitter (e.g., due to lower transmission rates, link layer retransmissions, use of long propagation delay links, or dynamics)
- lower physical security of media (e.g., due to lack of physical control over media)
- potential limited energy considerations (e.g., conservation of battery life)

In the past, mobile wireless networks were typically looked upon as homogeneous RF media problems. With time and the proliferation of numerous proprietary radio networks, the need for heterogeneous interoperability across networks is becoming a more prevalent interest area. The support of heterogeneity was one of the great successes demonstrated by Internet Protocol (IP) technology. In the near future, even commercial computing and network routing devices may typically have multiple wireless media interfaces (e.g., Bluetooth, 802.11, GSM). This network-

wide heterogeneous characteristic is especially true when considering military wireless systems (e.g., SATCOM, UHF LOS, other) and is often a driving design factor in building interoperable, adaptive information networks.

As mentioned, the mobile and wireless nature of a communication network has a profound influence on the layered design model. Below we use a simplified reference model to discuss these issues. This simplified model is more aligned with the logical layering realized within the Internet Protocol (IP) Suite.

Application Layer	Evolving multimedia applications Adaptive applications
Transport Layer	Application multiplexing End-to-end data reliability Congestion and flow control
Internetwork Layer	Heterogeneous Dynamic Routing Device Location Service Location Enhanced Queueing/Signaling/Policy
Subnetwork Interface/ Data Link Layer	Media Access and Control Broadcast, NBMA, point-to-point Multiplexing Subnetwork Routing/Bridging Link layer reliability and retransmission
Physical Layer	Modulation and Coding Antenna Characteristics (Omni vs. Directional)

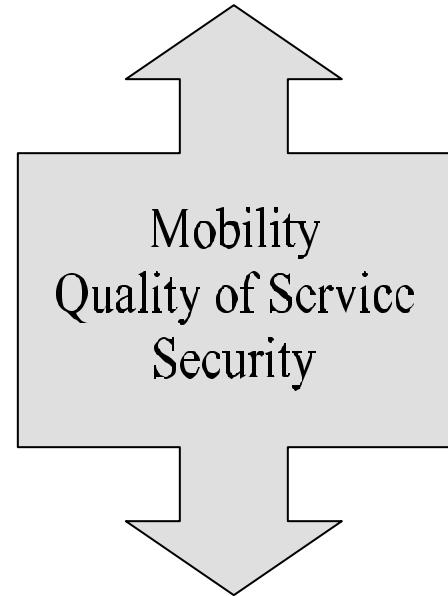


Table 1: Simple Layered Model and Functionality

Table 1 provides one example of viewing a simplified layered reference model and the potential functionality realized at different layers that may be important for consideration within a mobile, wireless network. In this paper, we largely concentrate on discussing the darker colored areas in the model with some mention and brief discussion of areas of lighter color.

We illustrate in Table 1 that mobility, security, and quality of service requirements effect and influence the design choices and engineering tradeoffs across all layers of the model. It is important to keep this in mind when reading discussions that follow. In many cases, this is a competing design space and engineering compromises need to be considered at a system level. In our subsequent discussions we maintain a network-centric perspective; thus we focus primarily on the transport and network layer issues relating to mobility and network dynamics. We will cover and discuss topics in the following general order:

- Heterogeneous Internetwork Routing
- End System Mobility
- Other Elements (e.g., Mobile Transport)

Heterogeneous Internetwork Routing

Routing technology provides multi-hop relaying and dynamic internetwork connection support. In mobile wireless networks, the performance characteristics of routing functions in the face of increased dynamics are of paramount interest. The ability to "self-configure" and "adapt to change" with a high degree of robustness is assumed as a fundamental requirement. In homogenous subnetworks, a routing function at the subnetwork layer may provide some limited multi-hop relaying and other services within that subnetwork domain. In heterogeneous networks, the IP internetwork layer interconnects different media segments and supports a variety of routing protocols to provide a media independent connectionless relaying function. We will here focus our discussion largely on the role of IP layer routing in the face of topology dynamics and mobile, wireless operation. Figure 2 illustrates an architecture where an IP layer mobile routing domain forms part of a larger heterogeneous internetwork and contrasts its functionality with that of other ad hoc wireless network technologies.

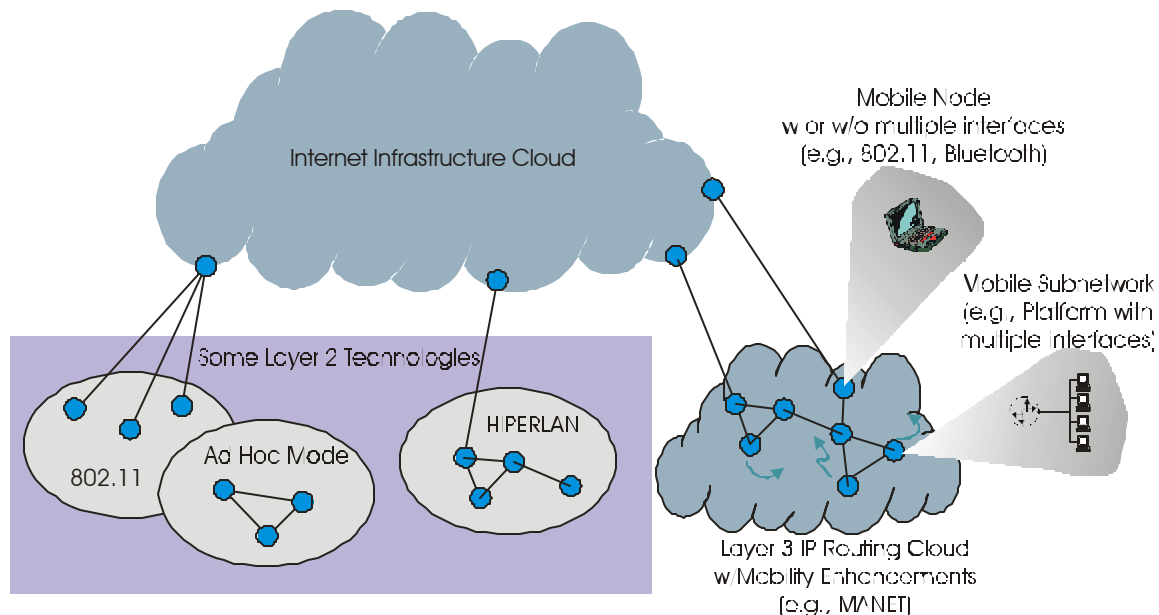


Figure 2: Mobile Ad Hoc Network Routing

The Office of Naval Research (ONR), the Defense Advanced Research Projects Agency (DARPA) and other DoD organizations have been supporting investigation and development of technology for both subnetwork and internetwork layer mobility over the past several years. While subnetwork layer mobile routing can solve a number of problems for limited mobile networking, higher layer "mobile adept" routing for internetwork interoperability is typically still required or it may also be used exclusively as the means for multi-hop connectivity within a network, especially when the network is heterogeneous. The IP protocol suite provides internetwork interoperability already; however, the present off-the-shelf IP routing protocols may not be well suited for military mobile networking architectures. As traditional IP routing protocols were not originally designed for the expected resource constraints and behavioral dynamics of the mobile wireless networking environment, there has subsequently been considerable research interest in the development of alternative routing solutions for inclusion as future IP technology standards.

The Internet Engineering Task Force (IETF), the primary standards body for Internet-based protocols and technology, has established a technical working group (WG) on Mobile Ad Hoc Networks (manet) to develop and evolve IP mobile routing protocol specifications and introduce them into the Internet Standards track. After some brief introduction and discussion of existing Internet routing technologies, we will examine in more detail the various technologies and proposals begin put forth within the IETF manet WG. Note that the IETF is essentially a standards organization; thus, much of the research and development of the various protocols under consideration has been supported by other research organizations and projects. For completeness, we discuss some relevant protocols that have not been proposed within the IETF manet WG, but provide a similar capability to those which have and should be considered.

Broad Internet Routing Architectural Considerations

A two level hierarchy is commonly used to describe IP routing in the Internet. In the case of unicast routing, *exterior gateway protocols* are used for routing between autonomous systems in the Internet backbone, while *interior gateway protocols* are used for routing within an autonomous system. The demands and requirements of these two routing functions differ significantly; thus, the preferred protocols for interior and exterior routing are typically not the same. A hierarchical architecture design based on defining autonomous systems and multiple routing domains is also applicable to broad scale military networking environments. Typically, an autonomous system comprises a set of networks and routers controlled by a single administrative authority. In military networks, autonomous systems may correspond to separate commands, branches of service, or allied forces. In addition to defining autonomous systems and routing domains based on administrative authority, architectural boundaries between portions of an internetwork with significantly differing networking environments should also be considered. One of these possible boundary conditions could be to determine the specific areas of significant mobility or dynamics vs. the more static areas. Defining multiple routing domains allows for the use of different protocols and services within different regions of an internetwork, even if they are under a common administrative authority.

Supporting end system mobility across subnetwork and routing domain boundaries in an internetwork requires additional mobility management solutions at higher layers. The architectural design and consideration of mobile routing and end system mobility management approaches should not necessarily be treated as orthogonal processes. As an example, the application of highly scalable routing solutions can relax range of mobility restrictions—reducing both the overhead associated with managing mobility across routing domains and the overall complexity of the system architecture. It is also possible to integrate end system mobility management with a mobile routing protocol, or develop techniques for end system mobility management that are based on mobile routing protocols. We will further discuss these possibilities and others in a subsequent section on end system mobility.

Some Commercially Available IP Routing Approaches

Within the fixed Internet, the Exterior Gateway Protocol (EGP) [Mills 84] and the Border Gateway Protocol (BGP) [RL 95] are the most commonly used exterior gateway protocols, and the Routing Information Protocol (RIP) [Hedrick 88, Malkin 98] and the Open Shortest Path First (OSPF) protocol [Moy 98] are the two most commonly used interior gateway protocols.

These commercially available protocols are designed primarily for operation in quasi-static hardwired networks. In such networks bandwidth is relatively abundant, and the protocol design goals involve minimization of router state (to permit scalability), low processing overhead, and reasonable reaction to topological changes when they occur (to ensure good packet routing performance). In other words, minimizing storage, computation, and time complexity are primary design goals of commercial Internet routing protocols intended for high-speed interface operation. Use of these protocols in wireless network environments should be carefully considered as the performance requirements and constraints may be quite different. As an example, due to more limited bandwidth operation the cost of additional processing burden may be a reasonably sound engineering tradeoff.

Exterior Gateway Protocols

EGP was the original exterior gateway protocol used for routing between autonomous systems in the Internet backbone. Its underlying mechanism is a distance-vector routing algorithm, based on the distributed Bellmann-Ford (BF) computation [BG 92]. However, the computation in EGP is based on metrics that represent a combination of preferences and policies; thus, limiting it to a reachability protocol rather than a shortest-path protocol. A significant limitation of EGP is that the routing backbone must possess a tree topology to prevent the formation of routing loops.

BGP is the preferred replacement for EGP. The underlying mechanism of BGP is a distance-vector routing algorithm, where the distance-vector entries are augmented with path and policy information. The addition of entire path information to the distance-vector entries prevents the formation of routing loops. An unusual difference in protocol signaling methods between BGP and other IP routing protocols is that BGP routers use TCP to communicate with each other. While this facilitates error management, routing updates are subject to TCP flow and congestion control behavior. The dynamics of this design choice when applied to wireless networks are complicated and not well understood at the present time.

For military network environments the use of EGP or BGP as an exterior gateway protocol may suffice provided that there is little or no mobility between autonomous systems. However, Even if this is the case, there are several areas for concern. An autonomous system with a mobile infrastructure may be more likely to partition. This can increase the dynamics and frequency of routing updates in the exterior gateway protocol. Also, for autonomous systems that have multiple boarder gateways, a partition may lead to inconsistent routing table entries. Note that BGP requires that all boarder gateways in an autonomous system communicate with each other (known as *internal peering*). A possible solution to this problem involves defining tunnels (between the boarder gateways) through adjacent hardwired autonomous systems to prevent partitioning between gateways. Finally, the use of TCP internal peering in BGP may be more problematic in mobile wireless networks due to the potential degraded performance of TCP in such environments.

Interior Gateway Protocols

RIP is based on a distance-vector routing algorithm and uses a hop count metric to compute the shortest path to a destination. Many distance-vector routing approaches suffer from the well-known “counting-to-infinity” problem that affects the convergence time of these protocols. There has been a large amount of theoretical and practical work seeking to remedy this problem

and improve the performance of distance-vector routing approaches. RIP uses techniques known as *split horizon* and *poison reverse* to combat the counting-to-infinity problem and speed the convergence time. However, these techniques do not completely address the problem. To further limit the slow convergence time associated with the counting-to-infinity problem, RIP uses a maximum hop count distance of 16 to represent infinity.

OSPF, a modern link-state routing approach, is the preferred interior gateway routing protocol on the Internet. In link-state routing approaches the entire network topology and the cost of each link are disseminated to all routers—allowing each router to independently compute paths to every destination. Dijkstra's shortest path algorithm [BG 92] is most commonly used for the local route computation. OSPF tends to be more operationally robust and has many features that make it more attractive than RIP for use in typical hardwire networks. In addition, OSPF adds the capability for type of service routing and load balancing to be implemented. OSPF also uses the notion of areas to limit the scope of topology dissemination and to increase the scalability over basic brute force link-state routing.

Both RIP and OSPF have design characteristics limiting their applicability in mobile wireless networks. In addition to the likelihood that these protocols will exhibit degraded performance in terms of increased control message overhead and convergence times in the context of mobile wireless networks, in some scenarios they may be essentially nonfunctional. This results primarily from differences in the assumptions regarding link-layer attributes, characteristics, and link layer interface configurations.

Both RIP and OSPF are designed based on a traditional IP routing paradigm, in that a router provides a forwarding capability between multiple networks to which it is attached. From the perspective of the IP routing layer, an attached network is a single hop, meaning that all the hosts and routers on the network can be reached directly using the lower-layer communication protocols of the particular network. However, for mobile wireless networks it is desirable to use IP routing to provide a multihop forwarding capability even over a single wireless communication technology. Specifically, given three wireless routers A, B and C (each with a single wireless communication interface), if A has connectivity with B and B has connectivity with C, but C does not have connectivity with A, then it is desirable to have B provide a forwarding capability between A and C. The standard specifications of RIP and OSPF are incompatible with this notion. For example, the split horizon mechanism in RIP prevents a router from propagating information about a route back over the interface from which the information was received. While OSPF supports three different types of interfaces (point-to-point, broadcast, or non-broadcast multiple access) and its behavior over each type of interface differs, none of the interface types can be used to support the desired multihop forwarding without resulting in efficiency or performance concerns. It is possible that OSPF extensions or link state variants could be developed and used for improved wireless and mobile network operation. In the subsequent section, we will discuss some of the recent advances and developments in this area.

Another significant limitation of RIP in particular, is the assumption of bi-directional connectivity between neighboring routers given the existence of unidirectional routes. That is, if router A receives a routing advertisement from router B, then it is assumed that router A can forward traffic through router B. For practical purposes in hardwired networks, this assumption generally holds true. However, in wireless networks, differences in transmission power levels,

local receiver noise, interference and propagation effects greatly increase the probability that unidirectional connectivity will exist between some router pairs. When using RIP, these unidirectional connections can result in the formation of dead-end routes or "routing black holes".

Emerging IP Mobile Routing Alternatives

We now present an overview and discussion of several mobile routing approaches, many of which are being worked on within the IETF manet WG. There are several interior gateway protocols that are specifically designed for mobile wireless networks. While such protocols are not yet widely available commercially, there are several prototype implementations in various stages of development and under commercial consideration. Some of the protocols are briefly discussed below. The discussions primarily focus on protocols with implementations that are deemed to be relatively mature, or that possess unique attributes of particular interest for discussion or debate.

While some of the mobile wireless routing approaches proactively maintain routes between all source-destination pairs, many have adopted a more reactive, on-demand design. Classification of a protocol into one of these categories is not a simple black or white decision, some of the protocols have elements of both or can be viewed as a hybrid design. Each protocol family has its advantages and disadvantages and appropriateness of the design type is affected by a spectrum of architecture and performance issues [CM 99].

In on-demand routing approaches, routes are reactively established to a given destination when needed (i.e., traffic driven). This design choice is based on the notion that in a dynamic topology it may not be necessary (or desirable) to maintain routes between all source-destination pairs at all times. The overhead expended to establish a route between a given source-destination pair will be wasted if the source does not require the route prior to its invalidation due to topological changes. The validity of this design decision is dependent in part on the traffic distribution and the topology dynamics in the network. Conceptually, it would seem most advantageous when traffic patterns are relatively sparse and topology dynamics are relatively high. While these approaches have the potential to reduce communication overhead, this is achieved at the expense of increased route acquisition latency.

The proactive approaches (also referred to as table-driven) are more similar in design to traditional IP routing protocols; thus, they are more likely to retain the behavior features of present routing protocols used in practice. Existing transport protocols and applications are more likely to operate as designed over proactive routing approaches than over on-demand routing approaches. In general, proactive routing approaches may be better suited than on-demand approaches when routes to a large percentage of the possible destinations are typically required. When network traffic patterns include a large percentage of possible source-destination pairs, the advantage of building routes to specific destination on-demand is reduced. In short the overhead of the route request process can be saved in situations where the demand can be assumed.

Several of the approaches include elements of both categories. For example, some of the predominantly on-demand approaches may cache known routes for future use or proactively maintain routes once they are initially established. In some approaches it is also possible to simultaneously support both on-demand routing for some destinations and proactive routing for

other destinations. This allows for improved robustness and route acquisition latency for frequently used destinations or services, while still benefiting from the overhead reduction for infrequently used destinations. Achieving the right balance between reactive and proactive operation may require some a priori knowledge of the networking environment or result in added complexity.

Ad Hoc On-Demand Distance Vector (AODV) Routing

The AODV protocol[PRD 00] is a source-initiated, on-demand, routing protocol. AODV protocol is designed to support both unicast and multicast routing. However, in this section we will limit the discussion to the unicast routing. While AODV shares some common attributes and design characteristics with other source-initiated, on-demand routing approaches being proposed within the manet working group, a primary difference is its use of destination sequence numbers to ensure loop freedom. This technique is an adaptation of the mechanism used in the Destination-Sequenced Distance Vector (DSDV) protocol.

AODV uses a query-reply mechanism for route discovery and building. In the case of AODV, either the destination or an intermediate node that has a valid route to the destination may reply to route request packets sent by a source. As the route request packet is propagated through the network, nodes that forward the request cache reverse routes to the source of the route request. A node that meets the criteria for replying to a route request for a given destination, unicasts a reply packet back to the source of the request. As the reply packet propagates back to the source, nodes that forward the reply establish route entries for the corresponding destination. Both route request and route reply packets include a hop count field that is used to compute the distance of routes along the path that the packets propagate. All route entries maintained in AODV, include an associated timer field that is used to expire and remove unused or stale routes.

For each route maintained by a node, the node also maintains a list of neighboring nodes that have recently used the route. These nodes are notified via a route error packet when the route is no longer valid. A route error packet may be triggered by detection of a broken link, reception of route error packet, or reception of data packet for which there is no valid route. The AODV specification includes a variety of options for sensing neighbor connectivity and detection of broken links. While discussion is limited here, the destination sequence number mechanism plays an important role in the processing of route request, reply and error packets and is instrumental in the prevention of routing loops.

At present, AODV provides limited support for subnet routes. Due to the need for destination specific sequence numbers, a subnet leader must be selected to generate the sequence numbers and send route replies for the subnet. Currently, there is no mechanism for dynamic election of a subnet leader, and this limits the practicality of the approach. AODV also defines a mechanism for controlling route request broadcasts, based on an expanding ring search using time-to-live (TTL) scoping. This technique is generally applicable to similar routing approaches. Mechanisms for enhanced quality of service, multicast routing, multihop broadcast, and address autoconfiguration have also evolved from the AODV specification. Recently, these specifications have been divided into separate documents. While the quality of service and multicast mechanisms are essentially integrated with the unicast routing protocol, the multihop

broadcast and address autoconfiguration mechanisms are completely orthogonal and thus generally applicable to other routing approaches.

While AODV's name implies it is a distance-vector algorithm, it does not provide a complete shortest-path computation. The distance values associated with routes in AODV are based upon the path over which control packets are forwarded, and does not guarantee discovery of the shortest path. The ability for intermediate nodes to respond to route requests is potentially advantageous when multiple sources are requesting routes to a common destination. However, if multiple sources rarely require routes to the same destination, a route request will typically need to propagate all the way to the destination for a reply to be generated. Thus, the network traffic patterns may significantly affect performance tradeoffs. AODV also includes many different timers that may affect performance in various networking environments and scenarios. There is little data regarding protocol performance sensitivity to timer settings or providing insight into tuning the settings for different networking environments.

AODV has been shown, via simulation, to perform well relative to other manet routing approaches in the context of a mobile network using IEEE 802.11 wireless technology. There are also cases in some studies where other approaches were shown to perform better. Unfortunately, it is difficult to draw strong conclusions based on limited simulation studies of an evolving technology. Simulation performance results are sometimes biased by the addition of important protocol enhancements (such as TTL scoping of route requests) to one protocol without providing equivalent enhancements to competing approaches. In many cases, such enhancements are applicable to a variety of routing approaches and can have a profound effect on performance in a given networking scenario. Additionally, the range of networking environmental characteristics and scenarios simulated is often narrow, limiting the general applicability of the results. Prior to any adoption of a protocol of this type, additional simulations should be performed with more accurate traffic model and network characteristics.

The AODV protocol was also recently put under more formal protocol analysis methods at the University of Pennsylvania and a routing loop problem was discovered. A fix for this problem was also recommended and has been worked into the most recent protocol draft specifications.

There is a prototype implementation of AODV for the Linux operating system. Scientists working for Ericsson, not the authors of the AODV specification, have developed this implementation. It is reported to be based on an early version of the specification, prior to November 1998, and thus does not include many of the newer important features. Currently, it appears that this implementation is not publicly available.

Associativity-Based Long-lived Routing (ABR) Protocol

ABR[Toh 99] is another source-initiated, on-demand, routing protocol and is based on a concept of establishing routes that are expected to be long-lived. The source-initiated, on-demand nature of ABR is a common attribute shared with some of the other protocols being proposed within the manet working group, but its bias towards the establishment of long-lived routes is a unique feature. A route is said to be long-lived if it remains valid over relatively long period of time; thus, during this time period, the sequence of routers that provide the forwarding between the source and the destination do not migrate beyond the connectivity range of their immediate upstream and downstream neighbors.

The ABR protocol consists of three phases: (a) route discovery, (b) route reconstruction, and (c) route deletion. As with most source-initiated, on-demand routing approaches, a query-reply mechanism is used for route discovery and building. In the case of ABR, a source initiates this process by flooding a query only when the source requires a route to the destination. As query packets are forwarded towards the destination by intermediate routers, they are augmented with path information and link metrics. The most unique of these link metrics is the associativity ticks, which provides a measure of the longevity of a router's links with its neighbors. Collection of the associativity ticks metric requires periodic beaconing between routers, which the current specification of ABR assumes to be provided by lower-layer protocols. Upon reception of one or more of the flooded query packets, the destination selects the best route and sends a reply packet back to the source. In ABR, reply packets are essentially unicast back to the source by reversing the route specified in the reply packet. This process results in the establishment of next-hop forwarding information in the intermediate routers to construct a single-path route between the source and the destination.

When an established route is segmented due to movement of source, destination, or any intermediate router, the route reconstruction phase is invoked. The route reconstruction phase includes mechanisms for deleting invalid portions of routes and repairing routes through the use of a localized query flood. The route deletion phase is initiated when a source no longer requires a route to the destination. Route deletion is accomplished via a flooding mechanism that directs all the intermediate nodes to remove the specific entry from their routing tables.

The attempt to construct more stable routes through the associativity ticks metric has the potential to reduce control overhead associated with route maintenance or reconstruction due to network topology dynamics. Yet, the protocol also relies heavily on the use of flooding for route discovery, reconstruction and deletion—which can result in a significant amount of control overhead. The control overhead due to flooding is driven by both topology dynamics and traffic patterns. Due to the complexity of the design tradeoffs it is difficult to assess the relative performance of ABR to other proposed approaches within the manet working group without a comprehensive simulation study. However, qualitative comparison with similar approaches may lend some insight.

Unlike similar proposed approaches within the manet working group, ABR neither caches multiple or unused routes, nor does it allow intermediate routers to respond to route queries. For this reason, it is perhaps the most connection-oriented approach within the manet working group. The benefits or detriments of this design choice are expected to be heavily dependent on the network traffic patterns. The design is likely best suited for supporting long-lived stream traffic between a relatively small number of source-destination pairs. If network traffic comprises primarily connectionless datagrams and short-lived streams between a relatively large number of source-destination pairs, more proactive, connectionless approaches may be better suited.

There is a prototype implementation of ABR for the Linux operating system. The implementation includes working wireless code and is currently being used for both field and laboratory testing on several laptops. ABR is patented, so official permission and license are required.

Dynamic Source Routing (DSR) Protocol

DSR[JM 96, BJM 99] is also a source-initiated, on-demand, routing protocol; however, its use of source routing distinguishes it from other similar approaches. When using source routing, each data packet to be routed carries in its header the complete ordered list of nodes through which the packet must pass; thus, the packets carry all of the necessary routing information. Another distinguishing factor is that link status sensing is largely based on hop-by-hop acknowledgment of data traffic and does not require any periodic transmission of control packets.

The DSR protocol comprises two primary functions: route discovery and route maintenance. Route discovery refers to the process that a source uses to obtain a source route to a destination, which is accomplished via a query-reply mechanism. During the route discovery process a route request packet is propagated out from the source. As the route request is forwarded, the sequence of nodes traversed is recorded in the packet by the forwarding nodes. Copies of the packet are propagated hop-by-hop until either the target destination is reached or another node that can supply a route to the target destination. All source routes learned by a node are stored in a local cache. The DSR specification proposes aggressive use of the route cache to reduce the frequency that route discovery must be initiated (i.e., nodes should cache routes learned through forwarding of request, reply and data packets, and if possible routes learned via promiscuous snooping of packet transmissions by other nodes).

Data packets are forwarded through the network based on a source route specified in the packet header. If a node is unable to forward a data packet to next hop indicated in the packet header, a route error packet is generated by the node and sent back to the original source of the data packet to indicate the link is broken. Each node that overhears or forwards the route error packet also removes the broken link from its route cache. If the node that generates the route error packet contains an alternate route to the destination in its route cache, it may modify the source route in the data packet header and continue forwarding it towards the destination.

There are several additional features and optimizations of DSR, some of which may be applicable to other approaches. DSR includes mechanisms for both TTL scoping of route requests and rate limiting of periodic route requests when a reply is not received. Both of these mechanisms are generally applicable to similar routing approaches. As with AODV, DSR also includes a mechanism for supporting a multihop network broadcast; however, in the case of DSR, the mechanism is integrated with the basic unicast routing protocol. While DSR does not include support for multicast routing, it does support the forwarding of multicast data packets via the network broadcast mechanism. The DSR specification also proposes a mechanism for suppressing redundant replies to a given request, thereby improving bandwidth efficiency and reducing probability of packet collision. The proposed mechanism is somewhat specific to DSR, but may be adaptable to similar routing approaches. Finally, DSR defines a framework for supporting enhanced QoS support. This framework allows sources to monitor the state of paths through the network, request promises along a given path, or request promises during route discovery. Additionally, the framework allows sources to reduce the frequency that the source route must be included in data packet headers by using a soft-state mechanism to identify a path through the network.

As with AODV, the ability for intermediate nodes to respond to route requests is potentially advantageous when multiple sources are requesting routes to a common destination. Thus, the network traffic patterns once again may significantly affect performance tradeoffs. A potential

advantage of DSR is the aggressive use of the route cache to limit the need for route requests. However, in highly dynamic topologies this may also lead to an increase in stale routing information and route error messages. The use of source routing presents a potential scalability limitation as the diameter of the network increases. Increasing network size and diameter impacts the size of both control and data packets. Route request, route reply, and data packets all carry hop-by-hop source routing information. The proposed QoS framework potentially reduces that impact of the source routes carried in data packets at the expense of the additional soft-state mechanism. The performance tradeoff of this modification is not well known.

DSR has been shown to perform well relative to other manet routing approaches in simulations of a mobile network using IEEE 802.11 wireless technology. As previously mentioned in the discussion of AODV, simulation studies are often limited in scope and the protocols being studied continue to evolve, so care should be taken in drawing any strong conclusions regarding the general performance of a specific protocol based on prior simulation studies. However, prior studies have shown encouraging results regarding the performance of DSR (without the QoS framework modifications) in some scenarios.

There is a prototype implementation of DSR for the FreeBSD operating system. The implementation has been used to construct a physical testbed for evaluation of performance in the field. This testbed has also been used for various demonstrations. The implementation code can be downloaded from the website for the Monarch Project at Carnegie Mellon University (CMU).

Optimized Link State Routing (OLSR) Protocol

OLSR[JMQ 00] is an adaptation of link-state routing that is designed for use in mobile ad hoc networks. Like traditional link-state routing approaches the protocol is proactive in nature and disseminates network topological information to support computation of the shortest path to all known destinations without regard to network traffic requirements. However, OLSR has two unique features that distinguish it from more traditional approaches and make it better suited for use in mobile wireless networks. First, OLSR does not disseminate complete topological information throughout the network. Only a subset of the network topology is disseminated, thus reducing the communication burden required to support the shortest-path routing computation. Second, OLSR uses a novel technique for flooding the network topology information that is more communication efficient than tradition methods.

The unique features of OLSR are based, in part, on the concept of selecting multipoint relays (MPRs). Each node designates a subset of its neighbors as its MPR set. A node selects its MPR set such that packets sent by the node and forwarded by the MPR set will be received by all nodes within two hops of the originating node. Sufficient information to compute the MPR sets is acquired via periodic exchange of hello packets with neighboring nodes. The MPR sets of all nodes in the network are collectively used to more efficiently flood the network topological information. Additionally, the flooded network topological information is also based on the MPR sets. Information regarding the MPR set selections is included in the periodic hello packets exchanged with neighbors, thus each node is aware of which neighbors have selected it as an MPR. Each node periodically advertises (i.e., floods throughout the network) the links between itself and the nodes that have selected it as an MPR. This information provided by each node collectively forms the partial topology database maintained by each node and is sufficient to

compute the shortest-path routes between all nodes in the network. Each link advertisement in a node's topology database has an associated holding time, upon expiration of which the link information is no longer valid and hence is removed from the database. While OLSR is primarily a soft-state approach, it also includes event triggered flooding of link state advertisements following the failure of links between a node and its MPR set. This improves the protocols convergence time for failures that invalidate prior routes.

OLSR specification includes support for "sleep mode" operation as a means of power conservation. When a node enters sleep mode it stops sending all periodic hello and topology advertisement packets and thus will no longer be used as an intermediate node for routing. In addition to simply ceasing its routing functions, a node entering sleep mode also negotiates with its neighbors to intercept and store data packets that are destined for the node entering sleep mode. Given that the node wakes up from sleep mode prior to the pre-negotiated sleep period, it may request delivery of the data packets stored by its neighbors. While support for power conservation is an important consideration for mobile ad hoc networks and the techniques employed by OLSR are readily adaptable to other routing approaches, unfortunately, they may be of only limited utility. Intermittent cessation of routing functions by nodes in the network can have a detrimental effect on the availability and correctness of routes. For some protocols it may also lead to an increase in routing control overhead. As for the data interception and storage function, it is unclear whether this would be desirable given the expected behavior of current transport-layer protocols and existing applications.

As a proactive routing approach, OLSR is potentially well suited when routes to a large percentage of the possible destinations are typically required. It can also be expected to provide a delivery service that more closely matches the behavior of current routing approaches used in hardwired networks. Recall that existing transport protocols and applications may operate more seamlessly over a proactive routing protocol without significant modification.

The MPR concept exploited in OLSR is intended to reduce the overhead associated with disseminating the topology, thus making it more scalable than traditional link-state routing approaches and potentially better suited for mobile ad hoc networks. However, the behavior MPR selection has not been well studied in the context of a dynamic topology. Therefore, the expected performance of the protocol cannot reliably be assessed without a comprehensive simulation study. Also note that the primarily soft-state approach used in OLSR to disseminate the topology information may not be the best approach for relatively large dynamic networks, especially when the dynamics across the network are not uniform, or packet loss rates are high. In such scenarios, an approach more biased toward event driven triggered updates with reliable flooding may provide better performance.

There is a MAC layer implementation of the protocols upon which OLSR is based (HIPERLAN). The MAC layer implementation is available for Linux, FreeBSD and Windows operating systems. Permission or license from the HIPERCOM project at INRIA is required. In addition, IP layer implementations of OLSR are also under development at INRIA.

Source Tree Adaptive Routing (STAR) Protocol

The STAR protocol[GS 99a, GS 99b] is also an adaptation of link-state routing in which only partial topology information is disseminated through the network. As with OLSR the

dissemination of only partial topology information reduces the communication burden relative to traditional link-state approaches and makes STAR better suited for use in mobile ad hoc networks. While both OLSR and STAR can be categorized as partial link-state approaches, the underlying algorithms upon which the protocols are based are significantly different. The partial topology information in STAR is based on neighbor exchange of spanning trees.

The current specification of STAR assumes the existence of an underlying protocol that provides neighbor discovery and link status sensing. While there is some ambiguity in the present draft specification, it appears that STAR assumes reliable, error-free delivery of control packets sent from a node to its neighbors. In STAR, each node communicates to its neighbors the source-rooted tree that it uses to reach every known destination. The aggregation of a node's adjacent links and the source trees reported by its neighbors constitute a partial topology graph that is used to generate its own source tree. STAR defines two distinct approaches that can be used to update the routing information, one that supports a shortest-path computation and one that sacrifices routing optimality to further reduce the communication burden associated with maintaining the topology information. Under the shortest-path approach, a node sends routing updates when its source tree changes (e.g., due to a link failure, link addition, or upon processing an update received from a neighbor). Under the less optimal approach, a node sends updates regarding its source tree only when it loses all paths to one or more destinations, when it detects a new destination, or when it determines that local changes to its source tree can potentially create long term routing loops.

STAR attempts to capture some of the benefit of on-demand approaches by including a mode that sacrifices the optimality of routes in favor of reducing routing control overhead. However, in both of its routing update modes of operation STAR continually computes routes to all known destinations without regard to network traffic requirements. Thus, STAR is likely best categorized as a proactive routing approach, even when operating in the less optimal routing mode. As a proactive routing approach, STAR is potentially well suited when routes to a large percentage of the possible destinations are typically required. It can also be expected to provide a delivery service that more closely matches the behavior of current routing approaches used in hardwired networks. Recall that existing transport protocols and applications may operate more seamlessly over a proactive routing protocol without significant modification.

Unlike OLSR, STAR is essentially a hard-state approach, and thus does not periodically re-advertise link state information. As previously mentioned a hard-state approach may be better suited for relatively large networks with variably dynamic topologies; however, the tradeoffs between soft and hard state approaches in dynamic networks have yet to be thoroughly studied and thus are not well understood. A potential disadvantage of a hard-state approach is the need for reliable delivery of control packets between neighboring routers.

STAR has been implemented in gated for the FreeBSD operating system. Note that the use of gated for the implementation may facilitate porting of STAR to other operating systems. The gated implementation of STAR has been demonstrated in testbed internetworks consisting of both wireless and hardwired links. The implementation is available to the manet community upon request.

Temporally-Ordered Routing Algorithm (TORA)

TORA[PC 97, PC 99] is a distributed routing protocol for multihop networks that is designed to minimize the communication overhead associated with adapting to network topological changes. The basic underlying algorithm is neither distance-vector nor link-state; it is member of a class referred to as link-reversal algorithms. A logically separate version of the algorithm is executed for each destination to which routing is required. The protocol simultaneously supports both source-initiated on-demand routing for some destinations and destination-initiated proactive routing for other destinations. In either case, the protocol builds a loop-free, multipath routing structure that is used for forwarding traffic to the given destination.

TORA is designed to work on top of lower layer mechanisms or protocols that provide the following basic services: neighbor discovery and link status sensing, reliable in-order control packet delivery, and security authentication. The Internet Manet Encapsulation Protocol (IMEP) is a companion protocol that is being developed with TORA and is designed to optionally provide these services. Note that the intent is for IMEP to serve as a common foundation over top of which a variety of different routing protocols can be used (e.g., TORA, STAR, or AODV). The specification of IMEP has expired; however, the protocol is still in active development and it is expected that the documentation will be updated shortly.

The routing structure in TORA is based on a concept of associating a “height” with each node in the network. Links between neighboring nodes are assigned a direction based on the relative height values of the two nodes (i.e., directed downstream from the higher node to the lower node). Collectively, the heights of the nodes and the directed links between the nodes forms a multipath routing structure where all downstream paths lead to the destination. Following some network topological changes (e.g., link failure) some directed paths may no longer lead to the destination. Only topological changes that result in this condition trigger an algorithmic reaction; thus, many topological changes require no reaction at all. When a reaction is triggered it proceeds as a sequence of link reversals (caused by the re-selection of node heights), which re-orient the routing structure such that all paths again lead to the destination. In portions of the network that become partitioned from the destination, nodes set their heights to “null” and their adjacent links become undirected; thus, erasing invalid routes. The routing structure is initially constructed by assigning non-null height values to previously null valued nodes. In its on-demand mode, this is accomplished via a query-reply mechanism. As with AODV and DSR, intermediate nodes may respond to route requests. However, unlike other on-demand approaches the route request in TORA is persistent and need not be periodically refreshed when a reply is not received. In its proactive mode, construction of the routing structure is accomplished by flooding a control packet from the destination. Once proactive mode is initiated by the destination, the routing structure will automatically reconstruct routes in previously partitioned sections of the network when they become reattached.

As with other on-demand routing approaches, the network traffic patterns will likely have a significant impact on performance tradeoffs. As with AODV and DSR, the ability for intermediate nodes to respond to route requests is potentially advantageous when multiple sources are requesting routes to a common destination. The flexibility of both on-demand and proactive routing behavior on a per destination basis in TORA is also potentially advantageous in some scenarios. This design simultaneously exploits performance advantages of using on-

demand routing behavior for infrequent destinations and proactive routing behavior for frequent destinations (such as servers or gateways).

The basic underlying link-reversal algorithm upon which TORA is based has been shown in simulations to be more scalable and communication efficient than link-state routing algorithms. However, simulations of TORA operating over IMEP in the context of a mobile network using IEEE 802.11 wireless technology illustrated some potential performance issues. While the network sizes used in the simulation study may not have been sufficiently large to demonstrate the benefits of TORA, they did provide important insight regarding the use of TORA in wireless networks with contention based MAC layers. The simulations clearly exhibited the disadvantage of routing approaches that require reliable in-order delivery of control packets and exposed some design limitations of IMEP. Note that this result may be applicable to a variety of routing approaches. Prompted by the results of that simulation study, IMEP has been significantly redesigned to improve its efficiency. Many of the concepts and design ideas that have emerged from research in reliable multicast transport protocols have been adapted and incorporated into IMEP. The relative performance and scalability of TORA with the newer IMEP design should be reassessed in the context of the relatively large mobile networks for which it was designed.

There is an implementation of both TORA and the newly redesigned IMEP for the Linux operating system. The implementation supports both wireless and hardwired links and is designed to operate with any type of IP interface device for which there is driver support. The implementation is being used for both experimentation and demonstration in heterogeneous wireless testbeds. Currently, source code distribution is limited, but a binary distribution is available upon request. There is also a commercial wireless router product based on TORA and IMEP that is under development by Nova Engineering Inc.

Wireless Routing Extensions for the Open Shortest Path First (OSPF) Protocol

As described in a previous section, OSPF is a link-state routing protocol designed for use in hardwired networks and thus is not particularly well suited for use in multihop wireless networks. In the standard OSPF specification there is no provision for the use of a limited scope broadcast interface (e.g., a wireless interface where the scope of a broadcast transmission may be limited in by range, interference or other effects). In the DARPA funded On-Board Switch (OBS) project, modifications and extensions to the OSPF protocol were developed to support such interfaces and enable OSPF to be used for routing in multihop wireless networks [BT 00].

Typically, dissemination of link-state updates and synchronization of the topology database does not occur between all pairs of neighboring OSPF routers that are physically connected via broadcast type media segment (e.g., Ethernet). For broadcast type media segments, a single router is elected as the Designated Router (DR), and all other routers on the media segment establish a logical adjacency with the DR. Link-state updates and database synchronization only occur between these adjacent routers, which significantly reduces the protocol overhead on broadcast segments. In a dynamic wireless networking environment, there is typically no one router with which all other routers can communicate directly; thus, election of valid DR is not always possible. The new interface type defined in the OBS modifications eliminated the DR election mechanism, causing adjacencies to form between all pairs of neighboring routers. While this allows the protocol to operate over wireless interfaces it potentially increases the overhead associated with maintaining the topology database.

Another function of the DR in the standard OSPF specification is to advertise a link-state update representing the broadcast media segment. Since the OBS modifications eliminated the DR election, modifications were also required to correctly represent the wireless links in the topology database. The OBS modifications require that a point-to-point link advertisement be used to represent each adjacency formed over the wireless interface. Again, while this allows the protocol to operate over wireless interfaces it potentially increases overhead.

The OBS project also developed several other modifications that enhanced the operation or performance of OSPF in wireless networks. The link-state advertisement format and routing computation were enhanced to allow synchronized replacement of routing tables by using predicted information. This enhancement provided for faster convergence times and less packet loss when routing changes could be anticipated (i.e., controlled shutdown of a router). Enhancements that made use of signal strength indications on wireless interfaces were also developed. Signal strength was used both to condition the formation of adjacencies between routers and to advertise the cost of the link for routing. These enhancements are generally applicable to other routing approaches.

While there remain concerns about the scalability of the routing approach demonstrated in OBS, the basic architecture was demonstrated successfully within a real, heterogeneous wireless network with both ground-based and airborne routing nodes. If the network routing nodes are small in number and bandwidth is moderate then an approach such as OBS may be quite effective. Alternative link-state variants such as OLSR and STAR may be equally viable and more scalable if sufficiently developed.

The wireless routing extensions of OSPF that were developed under the OBS project are implemented in gated for the FreeBSD operating system. However, there are potentially some intellectual property right claims associated portions of the development. Note that the use of gated for the implementation may facilitate porting of the OSPF extensions to other operating systems. Finally, several other protocols were modified under the OBS project (e.g. DHCP) and it is unclear if the enhanced OSPF protocol can be used in a standalone fashion without modification.

Other Developments

There are many other relevant protocols and development efforts that have not herein been covered in detail. One effort in particular is the Density and Asymmetry-Adaptive Wireless Network (DAWN) work done under the DARPA Global Mobility (GloMo) project. The DAWN project was lead by a team of scientists at BBN and they investigated and developed techniques for autonomous topology control, density adaptive routing, jammer evasion, and asymmetry/unidirectional accommodation. Another potential useful protocol for wireless networks is the Topology Broadcast based on Reverse-Path Forwarding (TBRPF) protocol. TBRPF is a link-state routing variant that was developed under the DARPA Small Unit Operations (SUO) project. Like OLSR and STAR, the design of TBRPF reduces the overhead associated with flooding of link-state topology information; however, unlike these other approaches TBRPF disseminates full topology information. An IETF Internet Draft on TBRPF has recently been submitted to the manet working group [BOT00]. Finally, there are several other protocols that were at one time documented as Internet Drafts within the manet working group but have since expired; e.g., the Zone Routing Protocol (ZRP), Cluster Based Routing

Protocol (CBRP), Core Extraction Distributed Ad hoc Routing (CEDAR), Relative Distance Micro-discovery Ad Hoc Routing (RDMAR).

Potential Application of IP Routing Solutions

We have presented an overview of several commercially available and emerging IP routing protocols and have discussed the basic features, strengths, and potential limitations of each. A complete analysis of each protocol across a set of potential architectures is complex due to the large design space and the performance tradeoffs one might consider. As technical guidance to future system analysis in this area, there are a number of network parameters that should be considered as primary performance drivers. Some of the more important anticipated system characteristics include the following:

- Required network scalability (e.g., number of routing nodes, number of end systems within an area)
- Typical link bandwidths of operation
- Percentage of application types and anticipated traffic models
- Degree of mobility and/or wireless link dynamics expected
- Degree of network density and heterogeneity expected

Network scalability is often a key factor to consider when applying candidate routing protocols to a particular scenario. In general, there may be many different system or resources constraints that limit the scalability of a protocol—e.g., available bandwidth or data storage capacity. Since the overhead of the protocol typically grows as some function of the number of nodes, it is often a limiting factor given bandwidth constraints. Similarly, protocol state also typically grows with the number of nodes, and thus can be a limiting factor given data storage constraints. However, these are not the only performance factors of interest, as the convergence time routing protocols may also be affected by scalability. The scalability of a protocol is not completely orthogonal to other issues; it is often complex and difficult to analyze, especially in the face of mobility and increasing topology dynamics. For example, a limiting factor such as the amount of protocol overhead may also increase with increasing topology dynamics. Design approaches for achieving greater scalability often involve a tradeoff with other performance measures. This can be seen in the differences between proactive and on-demand routing approaches. Proactive routing approaches continuously maintain routes between all source-destination pairs. On-demand routing approaches can be seen, at a basic level, as an attempt to reduce routing protocol overhead when only a subset of the network is actually communicating end-to-end at any given time. The reduction in routing protocol overhead is achieved at the expense of potentially higher route acquisition times. However, if the relatively sparse traffic pattern assumption is valid for a given scenario, then on-demand protocols may provide a more scalable routing solution.

Anticipated link capacities and the number of nodes sharing a common channel within a wireless network are important considerations prior to developing any particular architecture design. Commercially available Internet protocols often assume or are pre-configured for more optimal operation in higher link rate conditions (e.g., T1 and greater). This lesson is well demonstrated from the previous ONR Communication Systems Network Interoperability (CSNI) project in which OSPF was adopted. The protocol timers had to be tuned to provide a design compromise

between protocol reaction to changes and reduced overhead for application in limited bandwidth wireless networks. The lower link capacities and the increased dynamics characteristic of wireless networks begs for routing approaches featuring more efficient routing overhead while maintaining timely reaction to topology dynamics. This tradeoff space is a performance feature that forms a core design goal of most manet protocols under consideration.

Often overlooked in routing analysis, is the strong interaction with various traffic models presented to the routing protocols by the network applications and user/server distribution patterns. There are several areas of importance to consider. First, the number of sources and destinations and their relative distribution throughout the network is of importance. Second, the volume of traffic injected by sources and the long-lived nature vs. short-lived nature of interactions is important to consider. Each of these elements directly affects the potential performance and overhead efficiency of any routing protocol. As a simple example, if all nodes in a network are expected to be sources and destinations of unicast traffic at all times the efficiency gains expected from an on-demand protocol may be mute. In this case, a more proactive approach to routing probably makes more sense than an on-demand approach that tends to optimize for sparser source/receiver pairs. This consideration is also important when analyzing the multicast routing performance. In short, application characteristics and traffic distribution models should be included as a critical part of interpreting or understanding the applicability of any simulation or analysis results.

Finally, the degree and type of mobility or topology dynamics anticipated is another core performance analysis factor. Many of the emerging IP routing approaches previously discussed are designed to handle dynamics more efficiently. If little or no topology dynamics are anticipated in the topology then it is not necessary to consider a highly adaptive or efficient protocol as a solution. There will, however, be dynamics in any realistic wireless network whether there is considerable node motion or not. Some manet protocols are designed to be able to scale with the degree of motion much better than others. TORA is an algorithm that theoretically is designed to scale well with high levels of topology dynamics. Early ONR studies of TORA performance alongside ideal link state (ILS) demonstrated that TORA outperformed ILS in terms of overhead efficiency in scenarios with relatively high levels of dynamics and still performed reasonably close to ILS in scenarios with lower levels of dynamics. Other candidate manet protocols may work less well in high dynamics, but they may target more optimal route discovery given moderate dynamics. Simulation of these features need further study for a better understanding, but there are likely performance trends one can predict by understanding the nature and assumptions of individual algorithms.

As we have briefly discussed, the considerations for performance analysis of mobile routing protocols must take into account assumptions about various architectural design features. In some cases, multiple protocols may be chosen for use in different parts of a network for different reasons. This is perfectly acceptable and is not unlike the orthogonal autonomous domain designs that occur in present day wired networks.

Multicast Routing

The previous discussion of mobile-oriented and manet IP protocols focused on traditional unicast routing issues. Multicast routing is also an important area for consideration and discussion. We will keep the discussion on this subject relatively brief in this white paper as multicast mobile

routing technology is a relatively immature technology area and much of what is developed for unicast mobile routing--if proven effective--can be extended to develop multicast mobile routing variants. A number of multicast routing proposals for mobile environments have been presented and studied in recent years.

- Multicast AODV
- AMRIS
- CAMP
- LAM
- ODMRP
- Flooding Optimizations

Despite the relative immaturity of multicast routing technology for mobile environments, the rationale for adopting and using multicast technology is quite strong in wireless networks. The typical broadcast or semi-broadcast nature of ground mobile packet radio and satellite networks makes multicast dissemination of data intended for group communications a natural engineering choice. Full broadcast (every node is a receiver) and unicast (one node is a receiver) can be seen as subsets of a multicast delivery model.

End System Mobility

In IP-based internetworks, datagram routing information is largely based on networks and not individual hosts. The ability to aggregate and represent a contiguous set of individual host IP addresses by a single common network address and maintain routing information based on such network addresses is one of keys to the scalability of IP internetworking. However, the nature of this address aggregation also necessitates additional mechanisms to support the mobility of an end system (or host) across network, routing domain, or autonomous system boundaries. This end system mobility could refer to a laptop disconnecting from a wired network at the home office and reconnecting to a different network while on a business trip, or to a mobile wireless computing device dynamically associating with different points of attachment to a fixed infrastructure. In the most general case, an end system may be dynamically associated via a variety of different wired and wireless technologies to various points of attachment that are either part of a fixed or mobile infrastructure.

As stated in an earlier section, routing and end system mobility management are not necessarily orthogonal processes. The problem of supporting end user mobility can be largely solved via network layer routing techniques or higher layer location management techniques. Even if a higher layer location management solution is utilized it may be based on the existing routing technology or integrated with a particular routing approach. Performance assessment and comparison of techniques must consider the approach to both routing and end system mobility management, since protocol complexity, adaptability, efficiency, and communication overhead can be shifted between the two mechanisms based on the overall system design. In the remainder of this section we briefly discuss some of the existing and emerging technologies that may provide support for end system mobility.

Host Autoconfiguration

The Dynamic Host Configuration Protocol (DHCP) [Droms 97] allows for a host to dynamically acquire a temporary (or permanent) IP address and other network configuration information from a server upon startup or connection to a local network. The information provided via DHCP is sufficient to allow the host to communicate over a connected internetwork, and thus provides limited support for end system mobility. That is, each time a host is moved and connected to a new network the host may acquire a new IP address and other configuration information via DHCP, and subsequently send and receive IP datagrams using the new IP address.

Used alone, DHCP has two significant limitations in terms of supporting end system mobility. First, other hosts in the internetwork cannot locate or send datagrams to the mobile host without first learning the mobile host's new IP address via some other means. Second, connection-oriented data flows such as those using TCP will be terminated each time a mobile host changes its IP address. Dynamic Domain Name System (DDNS) updates [Mockapetris 87, VTRB 97] can be used to address the first of these issues; however, the second issue is a fundamental limitation of this architectural approach. When combining DHCP with DDNS, each time a mobile host changes its IP address via DHCP, a DNS update is sent. Thus, the mobile host's name in the DNS remains constant while its corresponding IP address changes with mobility. Other hosts in the internetwork may locate and acquire a mobile host's current IP address via a DNS request. Tracking frequent mobility changes of end systems via DNS updates has some potentially negative impacts on the DNS caching strategy and thus the efficiency of resolving DNS requests. There are security issues associated with endpoint authentication and key management that must be uniquely addressed when hosts are allowed to frequently change IP addresses.

Host Routing

Although IP-based datagram forwarding is primarily based on network routes, host-specific routes are possible. The use of host-specific routes provides another approach to supporting host mobility. In a host routing architectural approach, a mobile host maintains a constant IP address as it changes its point of attachment to the internetwork and dynamically updated host-specific routes are used to deliver datagrams to the host at its current location. Many of the routing approaches under consideration within the IETF manet working group include simultaneous support for both individual host and network masked routes. An advantage of this approach over the DHCP/DDNS architecture is that transparent upper layer connectivity to the mobile host can be maintained under dynamic conditions. That is, provided that the host-specific routing is sufficiently adaptive and converges relatively quickly, connection-oriented data flows like TCP need not be terminated when a mobile host changes its point of attachment.

Another example of this approach is the Local-Area Mobility (LAM) software [Cisco 00] available in the more recent Cisco Systems IOSes. A router configured for LAM monitors traffic on its network interfaces. When locally originated traffic from a host with an IP address that does not match the address and mask (or prefix) configured on the routers interface is detected, the router adds an ARP entry for the mobile host and installs a host route that points to the interface. The router may also redistribute the host route into any dynamic routing protocols being used in the internetwork, thus allowing the host route to propagate throughout the routing domain.

IP Mobility Support based on Encapsulation and Tunneling

The IP Mobility Support specification [Perkins 96a] defines a framework and protocols which provide tunnel based routing to mobile hosts. In the Mobile IP framework a mobile host is given a long-term IP address on a home network (i.e., its “home address”). When away from its home network, a “care-of address” that reflects the current point of attachment is associated with the mobile host. Routing to a mobile host that is away from its home network is supported via tunneling by mobility agents (i.e., a “home agent” on the mobile host’s home network and possibly a “foreign agent” at its current point of attachment). As the mobile host changes its point of attachment, it registers its current care of address with its home agent. Datagrams sent to mobile host’s home address are intercepted by the home agent and tunneled to the care of address via IP encapsulation. The end point of the tunnel (where decapsulation occurs) may be either at a foreign agent or at the mobile host. IP datagrams sent by the mobile host are using standard destination-based IP routing techniques.

As with host routing, Mobile IP supports transparent connectivity to the mobile host. That is provided that the mechanisms adapt and converge relatively quickly, connection-oriented data flows like TCP need not be terminated when a mobile host changes its point of attachment. A potential advantage of the Mobile IP approach is that it is transparent to the underlying routing approach; thus, it does not increase routing table size or in any other way impact the scalability of the underlying routing approach. The scalability of the Mobile IP architecture is primarily limited by the home and foreign agent resource constraints. There are several other issues regarding the Mobile IP approach that merit mention and consideration. First, the basic Mobile IP approach (for IPv4 internetworks) results in what is referred to a triangular routing. Datagrams sent from a corresponding host to the mobile host are initially sent to the home network, and then intercepted by the home agent and forwarded to the mobile host (in some cases this path may be very circuitous), while datagrams sent in the reverse direction are sent directly the corresponding host. Note that there is a present IETF Internet Draft addressing a potential solution to this issue [PJ 00]. Second, the architectural dependence on a home agent limits the robustness of the overall system design. A mobile host must be able to communicate with its home agent in order to register its care of address. Finally, for some applications the additional per packet overhead associated with IP encapsulation [Perkins 96b, Perkins 96c] may be of concern.

Many of the manet routing approaches mentioned can co-exist and support the overlying operation of Mobile IP techniques when and where sensible so the choice of mobile architecture protocols can be quite rich and should not be limited to protocol A vs. B, but rather what correct combination of overall mobile system components are sensible given the network operating conditions and application requirements.

Emerging End System Mobility Alternatives

A recent mobility management development worthy of some discussion is the approach used in the DARPA funded On-Board Switch (OBS) project. Architecturally the internetwork framework developed comprised a mobile wireless backbone infrastructure that employed a modified version of OSPF for routing and wireless end systems that dynamically attached to the mobile infrastructure supported by a modified version of Mobile IP. The end system mobility approach can conceptually be viewed as a more distributed variant of Mobile IP. Each backbone

node included both home and foreign agent functionality, and home to care of address binding information was disseminated among all backbone nodes. This approach eliminated the dependence of a mobile host on a single home agent; thus, increasing the robustness of the design. It also eliminated the use of triangular routing. However, it still used tunneling for routing datagrams between backbone nodes. The advantages described above were essentially achieved at the expense of flooding the address binding information, which significantly limits the scalability of the design. The design is also an example of an integrated approach, in that OSPF opaque LSAs were used to disseminate the address mobile IP binding information throughout the backbone. Thus, the routing protocol mechanisms also supported end system mobility and made enhanced mobile IP operation relatively seamless to the end systems involved. However, this makes the approach dependent on the use of OSPF (with support for opaque LSAs) as the underlying routing approach.

Another relevant development is the Mobile IP Router extensions to Mobile IP that were developed during a recent NC3A project. This development essentially extended the concept of tunneling to a mobile host to the case of tunneling to a mobile subnetwork. This provides a framework that could be used to support transparent connectivity to a subnetwork or portion of an internetwork that is aboard a mobile platform (e.g., a ship) and is dynamically changing its point of attachment. Performance tradeoffs of this approach should be further evaluated and compared to solutions based on the use of dynamic internetwork routing protocols.

Finally, within the IETF, private industry, and the research community there is considerable recent interest in supporting IP networking to handheld cellular end systems. The general consensus is that Mobile IP as currently specified is not sufficient to completely support this application. Within the IETF Mobile IP working group there is an Internet Draft that proposes extensions to Mobile IP to support operation in third generation cdma2000 networks [Xu 00]. There are also several IETF Internet Drafts that propose more general solutions based on an architecture where Mobile IP is used to support interdomain mobility and host routing is used to support intradomain mobility [OC 00, CGWKTV 00, RLTVS 00]. In this context, some of the manet routing solutions are being considered as scalable adaptive protocols for supporting the intradomain host routing. This work is likely to continue to evolve, but is indicative of the potential widespread applicability of the manet routing solutions.

Mobile Data Transport and Congestion Control

Initial research on Internet-type services in mobile wireless networks has primarily focused on network-layer issues—such as the design and development of routing protocols, tailored for operation in a highly-dynamic and bandwidth-constrained networking environment. While connectionless routing of packets is an important first step for providing Internet-type services, many applications require additional functionality such as end-to-end reliability or flow control. As depicted in Table 1, these end-to-end services are typically partially or fully provided by transport-layer protocols. Providing robust, functional transport-layer services in mobile wireless networks remains a largely unexplored research area.

As with existing network-layer protocols, traditional transport-layer protocols designed for use in hard-wired networks may not be well suited for use in mobile wireless networks without modifications or additional system component enhancements. The underlying assumptions used in the protocol designs may not be valid due to differences in the characteristics of the

networking environment and the services provided by lower-layer protocols. Ultimately, the transport layer performance would be less of an issue if the lower layer mobile, wireless protocols could provide a service that would closely emulate a fixed, wired network service. This totally mobile, wireless transparent design is not easily achieved and it is often better to reach some compromise design between lower and upper layer protocols for more adaptive performance. In short, research and investigation of transport and application layer issues in mobile wireless networks and further developments are essential for providing a complete set of Internet-type services comparable to those available in hard-wired networks.

Traditional Acknowledgment, Sliding Window Based Techniques

In the IP suite the TCP provides an end-to-end *reliable stream transport service*, which adds considerable functionality to the underlying best-effort, connectionless, IP service. TCP provides reliable, in-order delivery of a data stream from a source to a single destination—i.e., it provides reliability for unicast (one-to-one) data streams. The reliability mechanism of TCP is based on a technique of positive acknowledgment with retransmission and utilizes the concept of a sliding window to increase network utilization and throughput. The window size in TCP is variable and is dynamically adapted to control the injection of traffic into the network in response to congestion in the network or receiver buffer overflow. The mechanisms for adapting to congestion are largely based on an assumption that packet loss is due to congestion—an assumption that may not be valid in the context of a mobile wireless network. TCP also includes adaptive timer values, which are based on estimation of the mean and variance of round-trip-time delay. The techniques used to adjust these values can also have a significant effect on TCP performance. There has been considerable work in the area of tuning the performance of TCP for traditional hardwired networks. Furthermore, many extensions and modifications to TCP have been proposed (e.g., TCP Selective Acknowledgment) to improve the performance of TCP over connections with atypical characteristics (e.g., fixed satellite operation).

Negative Acknowledgement, Rate Based Techniques

As support for multicast (i.e., one-to-many or many-to-many) group communication is emerging and becoming more widely available, research on providing reliability for group communications is gaining interest. Emerging techniques for reliable multicast differ significantly from traditional acknowledgment and sliding window techniques. Achieving scalability with respect to the number of receivers in a multicast group, favors *negative acknowledgment* based reliability mechanisms and other mechanisms that limit feedback to the source. Initial reliable multicast approaches focused primarily on the development of robust, scalable and efficient designs and were *not* adaptive to network congestion or receiver buffer overflow. The rate of traffic injection into the network was typically a pre-configured or capped parameter.

The lack of effective congestion and flow control techniques for reliable multicast limits the practicality of the solutions for bulk transport over heterogeneous network conditions. Improper configuration of the rate of traffic injection can lead to poor utilization of network resources, congestion, or unfairness. A present widespread problem is that *non-adaptive* unicast or multicast traffic can cause other well-behaving TCP streams to unfairly slow down and potentially disrupt large number of connections as network resources become congested. A second related problem is that when congestion is absent, non-adaptive multicast traffic streams

cannot adjust to fully utilize the network resources or compete fairly, since conservative a priori rate limiting is often applied.

Overall, the problem of dynamically adapting to congestion is significantly more difficult for multicast when compared with unicast connections. The existence of multiple paths between data source and receivers and the potentially large scale in number of receivers complicates any control function that estimates, senses, and reacts to dynamic congestion within a network. This problem is likely further compounded in mobile wireless networks by more frequent capacity fluctuations and/or dynamic routes that introduce additional transients in the system.

Recent advancements have led to the development of *rate-adaptive* techniques for congestion and flow control of reliable multicast streams. The approach dynamically adjusts the *rate* of traffic injection into the network based on congestion indicators. This is in contrast to traditional techniques that control traffic injection based on a sliding window. The techniques are still somewhat experimental, but initial performance results are promising. To date, related work has not focused on operation in mobile wireless networks. In fact, the focus has been on the development of techniques that achieve fairness with TCP streams—in essence, mimicking the behavior of TCP. A better understanding of these approaches and their potential performance in mobile, wireless scenarios is needed.

An Approach for Mobile Wireless Networks

There are several expected characteristics of mobile wireless networks that limit the viability of prior techniques. Primarily, a significant amount of packet loss may *not* be due to congestion. Thus, the validity of a fundamental assumption upon which prior techniques are based must be challenged. Secondly, the dynamics of a mobile network will likely result in more frequent changes in both bandwidth and routes between source-receiver pairs. This introduces additional transients in the system and will likely also lead to an increase in out-of-order delivery and delay variance. All of these aspects can have a significant impact on the performance of congestion control techniques.

The use of network assisted congestion indicators—e.g., Explicit Congestion Notification (ECN) or source quench—may provide an improved capability for early sensing and reaction to network congestion in wireless networks. With techniques such as ECN, packets are tagged with congestion information as they are forwarded through the network. This provides a potentially less ambiguous indication of congestion than sensing packet loss at the transport protocol layer in the mobile wireless networking environment. However, ECN tagged packets may be subsequently lost prior to reaching the receiver; thus, end-to-end transport layer measures of delay and loss are still important indicators. A *hybrid* of both *network-assisted* and *end-to-end* congestion indicators will likely provide the best compromise solution. The challenge is in determining when and how to tag packets with congestion information and how to combine the network-assisted indicators with end-to-end indicators to improve performance over traditional techniques. The *ideal* solution would be a unified approach that is applicable to both unicast and multicast data streams, provides good network utilization, limits loss due to congestion, works with existing or planned infrastructure components, and provides fairness between competing streams.

Mobile Applications and Data Services

While we do not discuss mobile applications and services within this paper, it is the authors' opinion that this is an important design and consideration area in an overall mobile architecture. Resilient and adaptive applications that can continue to perform effectively under degraded conditions can significantly enhance network operations from a user's perspective. Such applications can also ease the design pressure significantly in complex engineering areas such as QoS and mobile routing at the network layer.

Application and data services that also better manage connections and are less "state and location dependent" will likely perform better and allow more useful information management within a mobile network scenario. We recommend more consideration and design in these areas in concert with the overall architecture and the transport and network layer issues laid out earlier in this document.

Summary

We have presented a brushstroke overview of mobility-enhanced internetworking protocols and their applicability to mobile and dynamic wireless communication architectures. Our focus is on networking technology to provide heterogeneous support and interoperability with existing and planned IP-based infrastructures and applications. The intent of this paper is not to provide detailed architectural answers to a specific system design, but to provide initial guidance for further examination of architectural and protocol tradeoffs in near term R&D networking projects (e.g., ONR INSC). We have outlined engineering tradeoffs, technology status, and architectural options surrounding a number of areas including: mobile ad hoc networking (manet) routing and related evolving standards work, mobile host and router support issues (e.g., Mobile IP), multicasting, mobile QoS, congestion control and transport layer performance. This paper provides the initial guidance in performing a more detailed system analysis and design effort for a complete mobile architecture. We wish to reiterate that a well-designed, mobile network architecture is likely made up of many combinations of protocol components, a broad set of which we have discussed. We conclude this paper with a number of considerations to be addressed prior to a design and protocol selection effort for a specific wireless, mobile architecture.

- How mobile or dynamic is the network likely to be (how frequent are the topological dynamics, the wireless link perturbations)?
- Is there mostly edge system mobility or is the infrastructure itself dynamic? Is it likely some combination?
- How heterogeneous is the wireless internetwork (number of RF media and link types)?
- Will the system support high bandwidth or mostly constrained links?
- What is the general form and distribution of network applications to be supported?
 - What is the expected pattern and load of network traffic driven by the applications?
 - What QoS requirements are driven by the applications?

- Is the architectural design constrained or influenced by administrative, existing infrastructure, or system capability considerations?
 - Is there a natural or desired backbone?
 - Are there natural or desired routing domain boundaries?
- Is there a network node autoconfiguration requirement?
- What are the security requirements and how is security managed?
- Are the systems power-constrained (e.g., portable batteries)?
- Are multicast services required/desired?

Armed with the full or partial answers to these questions, a high-level engineering tradeoff analysis can be performed to set further direction and detailed focus on mobile system design efforts. The issues and technical overviews we have provided in this paper will assist in that engineering process.

References

- [BG 92] D. Bertsekas and R. Gallager, *Data Networks*, Prentice-Hall, 1992.
- [BJM 99] J. Broch, D. Johnson and D. Maltz, "The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks," *IETF Internet-Draft draft-ietf-manet-dsr-03.txt*, October 1999.
- [BOT 00] B. Bellur, R. Ogier and F. Templin, "Topology Broadcast based on Reverse-Path Forwarding (TBRPF)," *IETF Internet-Draft draft-ietf-manet-tbrpf-00.txt*, August 2000.
- [BT 00] BBN Technologies and TRW Tactical Systems, "Software Design Document for the On Board Switch CDRL No. A007," *Document No. D2-8346*, U.S. Air Force AFMC, Rome Laboratory, January 2000.
- [CGWKTV 00] A. Campbell, J. Gomez, C-Y. Wan, S. Kim, Z. Turanyi and A. Valko, "Cellular IP," *IETF Internet-Draft draft-ietf-mobileip-cellularip-00.txt*, January 2000.
- [Cisco 00] Cisco Systems, "Cisco IOS Local-Area Mobility—A Cisco IOS Software Solution to Business Needs to Enable Mobility within the Enterprise Network," *White Paper* http://www.cisco.com/warp/public/cc/pd/iosw/ioft/lam/tech/lamso_wp.htm, Cisco Systems, April 2000.
- [CM 99] S. Corson and J. Macker, "Mobile Ad Hoc Networking (MANET)," *IETF RFC 2501*, January 1999.
- [Droms 97] R Droms, "Dynamic Host Configuration Protocol," *IETF RFC 1541*, March 1997.
- [GS 99a] J. Garcia-Luna-Aceves and M. Spohn, "Source Tree Routing in Wireless Networks," *Proc. of IEEE ICNP '99*, Toronto, Canada, October 1999.
- [GS 99b] J. Garcia-Luna and M. Spohn, "Source Tree Adaptive Routing (STAR) Protocol," *IETF Internet-Draft draft-ietf-manet-star-00.txt*, October 1999.
- [Hedrick 88] C. Hedrick, "Routing Information Protocol," *IETF STD 0034*, June 1988.

- [JM 96] D. Johnson and D. Maltz, "Dynamic Source Routing in Ad Hoc Wireless Networks," *Mobile Computing*, T. Imielinski and H. Korth, eds., Kluwer Academic Publishers, 1996.
- [JMQ 00] P. Jacquet, P. Muhlethaler, A. Qayyum et al., "Optimized Link State Routing Protocol," *IETF Internet-Draft draft-ietf-manet-olsr-02.txt*, July 2000.
- [Malkin 98] G Malkin, "RIP Version 2," *IETF STD 0056*, November 1998.
- [Mills 84] D. Mills, "Exterior Gateway Protocol," *IETF STD 0018*, April 1984.
- [Mockapetris 87] P.V. Mockapetris, "Domain names – implementation and specification," *IETF STD 0013*, November 1987.
- [Moy 98] J. Moy, "OSPF Version 2," *IETF RFC 2328*, April 1998.
- [OC 00] A. O'Neill and S. Corson, "Edge Mobility Architecture," *IETF Internet-Draft draft-oneill-ema-02.txt*, July 2000.
- [Perkins 96a] C. Perkins, "IP Mobility Support," *IETF RFC 2002*, October 1996.
- [Perkins 96b] C. Perkins, "IP Encapsulation within IP," *IETF RFC 2003*, October 1996.
- [Perkins 96c] C. Perkins, "Minimal Encapsulation within IP," *IETF RFC 2004*, October 1996.
- [PC 97] V. Park and M.S. Corson, "A Highly Distributed Routing Algorithm for Mobile Wireless Networks," *Proc. of IEEE INFOCOM '97*, Kobe, Japan, April 1997.
- [PC 99] V. Park and M.S. Corson, "Temporally-Ordered Routing Algorithm (TORA) Version 1 Functional Specification," *IETF Internet-Draft draft-ietf-manet-tora-spec-02.txt*, October 1999.
- [PJ 00] C. Perkins and D. Johnson, "Route Optimization in Mobile IP," *IETF Internet-Draft draft-ietf-mobileip-optim-09.txt*, February 2000.
- [PRD 00] C. Perkins, E. Royer and S. Das, "Ad Hoc On Demand Distance Vector (AODV) Routing," *IETF Internet-Draft draft-ietf-manet-aodv-06.txt*, July 2000.
- [RL 95] Y. Rekhter and T. Li, "A Border Gateway Protocol 4 (BGP-4)," *IETF RFC 1771*, March 1995.
- [RLTVS 00] R. Ramjee, T. La Porta, S. Thuel, K. Varadhan and L. Salgarelli, "IP Micro-mobility Support using HAWAII," *IETF Internet-Draft draft-ietf-mobileip-hawaii-01.txtt*, July 2000.
- [Toh 99] C. Toh, "Long-lived Ad Hoc Routing based on the Concept of Associativity," *IETF Internet-Draft draft-ietf-manet-longlived-adhoc-routing-00.txt*, March 1999.
- [VTRB 97] P. Vixie, Ed., S. Thomson, Y. Rekhter and J. Bound, "Dynamic Updates in the Domain Name System (DNS UPDATE)," *IETF RFC 1771*, April 1997.
- [Xu 00] Y. Xu et al., "Mobile IP Based Micro Mobility Management Protocol in the Third Generation Wireless Network," *IETF Internet-Draft draft-ietf-mobileip-3gwireless-ext-04.txt*, June 2000.